Tracking Performance at the Crossroads of Perception and Memory

A Senior Honors Thesis

Presented in partial fulfillment of the requirements for graduation *with honors research distinction* in Psychology in the undergraduate colleges of The Ohio State University

Emily Ruth Weichart

The Ohio State University
April 2013

Project Advisor: Dr. Per B. Sederberg, Department of Psychology

Abstract

One of the most robust effects in memory is the spacing effect--memory is improved for repeated stimuli separated by intervening items (spaced encoding) relative to immediate stimulus repetitions (massed encoding). Because there are individual differences in working memory capacity (the number of items a person is able to attend to simultaneously) we sought to define the relationship between the spacing effect and working memory. We worked under the hypothesis that the spacing effect works through a mechanism known as repetition attenuation—this is a reduction in processing for repeated items. When other items intervene in between repetitions, the amount of processing approaches baseline as a function of spacing. Studies have shown that as the combined amount of processing for the item presentations during the test phase increases, recognition memory improves. To test the theory that the spacing effect is due to differences in perceptual processing of repeated items, we designed a novel experiment that includes three spaced conditions and a massed condition. Famous and unfamiliar human faces were used as stimuli. We administered a measure of working memory capacity called the "Ospan" (operation span), and we predicted a direct relationship between o-span and memory performance, such that individuals with larger working memory capacities would show the greatest memory performance in the 'long' spacing condition. We found significant main effects of spacing and stimulus familiarity for recognition performance and study phase priming, which is a correlate of repetition attenuation.

Tracking Performance at the Crossroads of Perception and Memory

Picture this: you wander into the Berry Café for lunch, and you strike up a conversation with a kindly, curly-haired stranger while you are waiting in line. Once you have received your sandwich and are ready to leave, you wave a quick farewell to your mysterious new acquaintance. Two days later, you return to the Berry Café for coffee. Sitting in the library enjoying your beverage, something catches your eye: it is the curly-headed visage of the very same stranger you spoke to mere days ago. Instantly, the memories of your previous encounter flood your mind's eye. What you are experiencing at this moment is *recognition*. For recognition to occur, you not only need to see this elusive stranger more than once, but you also need to remember the event of seeing him for the first time.

In an experimental setting, we can study recognition via repeated images. The phenomenon of change that occurs during re-exposure to a stimulus is termed *repetition priming,* and can manifest as differences in accuracy or reaction times. According to Bruce and Young (1986), repetition priming for faces results from the strengthening of the connections between 'face recognition units' (FRUs) and 'personal identity nodes' (PINs)—in other words, repetition priming reflects an ability to match what we perceive about a person to our internal database of who we have seen before. *Repetition attenuation (RA)* is a peculiar derivative of repetition priming, whereby there is a decreased neural response in cortical regions associated with processing the given stimulus when it is repeated (Grill-Spector, Henson, & Martin, 2006). As shown by recent neuroimaging studies, the manifestation of this effect appears to be intimately tied to the familiarity of the stimulus.

In a study by Henson, Shallice, and Dolan (2000), the experimenters used functional magnetic resonance imaging (fMRI) to record event-related hemodynamic responses to images of familiar (famous) and unfamiliar faces that were presented two or five times. In the bilateral fusiform cortex, they found an enhancement effect for subsequent presentations of the unfamiliar stimuli and an attenuation effect for subsequent presentations of the familiar stimuli. This means that when the stimuli were unfamiliar faces, there was more neural activation for the second presentation of a face in comparison to the first. When the stimuli were familiar faces, there was instead a reduction in neural activation. The changes were localized to a brain region that has been specifically linked to the processing of faces (Kanwisher, McDermott & Chun, 1997). Furthermore, the respective enhancement and attenuation effects became more pronounced with each additional presentation of the stimuli. In a follow-up study, Henson and colleagues attributed these findings to differences in perceptual representations for familiar and unfamiliar stimuli (Henson, Price, Rugg, Turner & Friston, 2002). When we see an unfamiliar face, we must form a representation of it so that we are able to recognize it later. For familiar faces, we can simply activate an existing representation, and match what we see to what we have already seen before. Hence, when unfamiliar faces are presented multiple times, the increase in neural responses may occur because of the additional recognition processes that are linked to forming new perceptual representations. The first presentation of a familiar face, however, activates an existing representation so that additional presentations require lower levels of processing.

What does this mean in terms of memory? The literature has yet to offer a consistent answer. When Henson et al. (2002) tested participants for their recognition memory of items seen in the study phase, the results were staggering: the hit rate versus the false alarm rate for identifying second presentations of familiar faces was 0.81 compared to 0.46 for unfamiliar faces. Since recognition performance for familiar faces was significantly better than performance for unfamiliar faces, it could indicate that repetition attenuation is associated with more efficient memory storage. However, categorizing an image as "familiar" indicates an overall higher instance of exposure (including contexts outside the confines of an individual experiment) and more elaborate representation-formation, in contrast to the images deemed "unfamiliar." For example, we have all seen thousands of pictures and videos of President Barack Obama. When most Americans look at a picture of his face, they can reliably recall not only his name and occupation but also their pre-established opinions of his personality and career. In reference to Bruce and Young's model (1986), repetition attenuation is a function of how readily available an existing representation is. When more processing has been collectively devoted to a particular stimulus, recognition will occur at a faster rate. It is not repetition attenuation itself that led to Henson's results (2002), but rather the divergent strengths of the perceptual representations.

In the present study, one of our aims was to find out the effect on memory when repetition attenuation is isolated from the strength of individual perceptual representations. In order to dissociate these two forces, we investigated the interaction between repetition priming effects and the *spacing effect*—the phenomenon whereby people are better able to remember items that are studied multiple times over spaced intervals relative to massed intervals (Cepeda, Pashler, Vul, Wixted & Rohrer, 2006; Donovan & Radosevich, 1999). A study by Xue et al. found a result of repetition attenuation in the bilateral fusiform gyrus modulated by spacing for unfamiliar faces: participants saw each of the face stimuli four times, either in the massed condition or randomly spaced with a mean lag of 12 items in between (Xue, Mei, Chen, Lu, Poldrack & Dong, 2011). As expected, they found that recognition performance on a memory test was significantly better for items learned under the spaced condition than the massed condition. Importantly, fMRI data showed that spaced learning significantly reduced repetition attenuation and that remembered faces elicited less repetition attenuation than forgotten ones in both conditions. In this case, spaced learning led to less repetition attenuation, which led to enhanced memory. Similar fMRI results were found when lag was varied between presentations of novel and repeated words (Callan & Schweighofer, 2009; Wagner, Koutstaal, Maril, Schacter & Buckner, 2000). These studies point to a counterintuitive possibility: priming actually hinders the efficacy of episodic encoding.

Researchers have proposed a variety of explanations for the spacing effect, each one accounting for different combinations of experimental evidence. Callan and Schweighofer's (2009) study with word repetitions resulted in the active rehearsal hypothesis, which suggests spaced repetitions (relative to massed repetitions) create a greater need to re-encode the stimuli and engage in more rehearsal. Though it is not possible to use subverbal rehearsal to remember a series of faces as participants may have done in Callan and Schweighofer's word study, there is still something to be derived from this hypothesis: when viewing a series of images, we devote attention and neural processing to each individual image as it appears. When an image is immediately repeated, it is possible that we bind the repetitions into a single event. This idea is related to the study-phase retrieval theory, which states that when we see an item for the second time, we automatically retrieve the memory of the first presentation of that item—massed repetitions do not offer that advantage (Greene, 1989). Hence, the process of encoding involves not only the individual stimulus, but the context as well; multiple contexts ensures that there are multiple cues available for retrieval.

The short-term perceptual priming hypothesis posits that the lack of intervening items in the massed condition leads to reduced perceptual processing of the second item presentation. This means that the sum total amount of processing devoted to a unique item is larger for spaced items (Russo, Parkin, Taylor & Wilks, 1998). Supported by robust findings that stronger brain activity during encoding predicts better subsequent memory (Brewer, Zhao, Desmond, Glover & Gabrieli, 1998; Wagner, Schacter, Rotte, Koutstaal, Maril, Dale, Rosen & Buckner, 1998), this hypothesis suggests that spaced learning therefore leads to improved performance on tests of recognition memory. The short-term perceptual priming hypothesis is somewhat related to the neural model of the spacing effect developed by Wagner et al. (2000): encoding itself involves directing attention to the features of the stimulus that are relevant to the task at hand (in the current experiment, participants are asked to make a male/female judgment for individual pictures of faces). Attending to task-relevant features results in perceptual priming, such that those features are readily available when the stimulus is seen for a second time. Hence, priming biases re-encoding such that task-irrelevant features receive less attention. These supposedly irrelevant features, however, have the potential to serve as additional cues for later recognition. Wagner and colleagues propose that spacing decreases the likelihood that individuals will rely on priming to complete the task at hand.

Though studies have investigated repetition attenuation in the context of stimulus familiarity (Henson et al., 2000; 2002) and as a function of the spacing effect (Wagner et al., 2000; Xue et al., 2011), no study to date has investigated how these two forces interact to affect recognition memory. The present behavioral studies involved manipulations of stimulus familiarity (famous vs. non-famous faces) and spacing (massed, short, medium, and long lag conditions). By measuring recognition memory for study items, we predicted that spacing would further intensify the repetition enhancement effect that Henson and colleagues (2000) found for presentations of unfamiliar faces. Also in accordance with Henson et al.'s findings, we predict an overall attenuation effect for repetitions of famous faces. To account for Wagner and colleagues'

(2000) hypothesis that priming effects are mediated by the spacing effect, however, we predicted that recognition memory would improve as a function of lag.

As an additional point of comparison, we measured individual differences in working memory capacity. Working memory (WM) is defined as an executive function serving to monitor the acquisition of new information and the retrieval of old (Baddeley & Hitch, 1974). We measured it via a modified version of the *operation span task (Ospan)* (Turner & Engle, 1989; automated by Unsworth & Heitz et al., 2005), which requires individuals to solve math problems while also remembering a string of letters. Several WM span tasks have utilized similar procedures: to-be-remembered items are interspersed with a distraction task, and the target items must be reproduced in serial recall (reading span: Daneman & Carpenter, 1980; counting span: Case, Kurland, & Goldberg, 1982; spatial span: Shah & Miyake, 1996). Such tasks are used widely in memory research, and are considered reliable and valid measurements of WM span (Engle, Tuholski, Laughlin & Conway, 1999).

We believe that there could be a relationship between WM span and the spacing effect—in a spaced retrieval task, items are less likely to be successfully remembered when retrieval events occur outside of an individual's working memory capacity. According to study-phase retrieval theorists (i.e. Greene, 1989), the spacing effect depends on successful retrieval of the original encoding event at the time of the second presentation of an item. If the first item has been forgotten by the time it is repeated, the benefit of spacing will be lost. To illustrate this, a study by Maddox and colleagues showed that older adults benefited most from very short spacing intervals (i.e., one intervening item) in a memory retrieval task, whereas younger adults benefited from longer spacing intervals (i.e., two or five intervening items) (Maddox, Balota, Coane & Duchek, 2011). The authors attributed this effect to differences in working memory capacity, which has been shown to deteriorate with age (McCabe, Roediger, McDaniel, Balota & Hambrick, 2010; Park, Smith, Lautenschlager, Earles, Frieske, Zwahr & Gaines, 1996). Applying this to the current study, we predict that individuals with larger WM capacities (as measured by the Ospan) will benefit more from longer lags of spacing than individuals with shorter WM capacities in terms of recognition memory performance.

Method: Experiment 1

The primary aim of our first study was to reconcile the divergent conclusions that research has presented about the relationship between spaced repetitions of unfamiliar faces and recognition memory. As aforementioned, Henson and colleagues (2000) found a repetition suppression effect when famous faces were repeated, but found a repetition enhancement effect for unfamiliar faces. This seems to contradict the findings of repetition attenuation observed via fMRI that Xue and colleagues (2011) found when they repeated images of unfamiliar faces at various lags--there was an effect of repetition attenuation across all conditions that correlated with a subsequent benefit of recognition memory. In addition, we wanted to see if memory for unfamiliar faces within conditions were dissociable by measures of working memory capacity.

Participants

For this study, 62 volunteers were recruited from the Research Experience Program (REP) for Introductory Psychology 1100 students at Ohio State University. All participants had normal or corrected vision, fluency in English, and were at least 18 years of age (M=18.7, SD=1.1). Participants were compensated for their one-hour time commitment with partial course credit. In concordance with standard Institutional Review Board (IRB) protocol, each participant provided consent before beginning the experiment.

Design and Materials

Stimuli were color images (480 x 480 pixels) of individual faces that included gender-distinguishing characteristics such as hair, makeup, and jewelry. Across all study and test blocks, a total of 320 unique images were presented. The images were 18 to 50-year-old unfamiliar faces obtained from an online database developed by Minear and Park (2004). Equal numbers of male and female faces were presented. This experiment consisted of four interleaved conditions, which were presented in a within-subject design so that each participant was exposed to every condition. **Figure A** below shows an illustration of the stimuli and conditions.

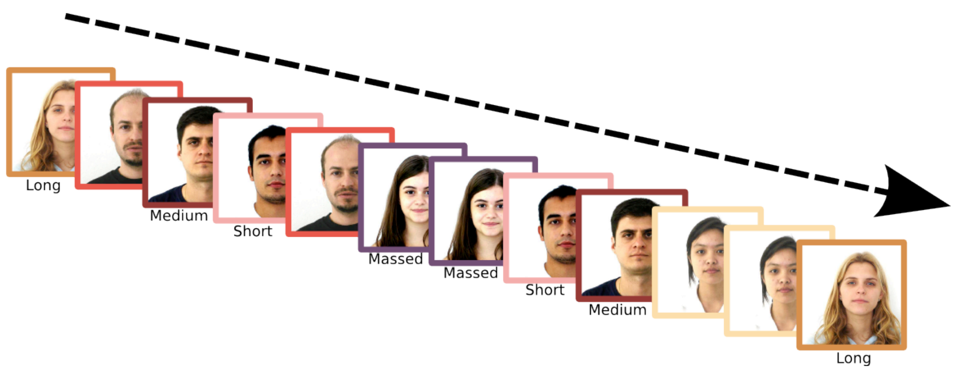

Figure A: Illustration of interleaved spacing conditions during the study phase of Experiment 1

In the study block, each stimulus was presented twice with four conditions of spacing in between: the massed condition had 0 intervening items, the short condition had 2-3 intervening items, the medium condition had 5-7 intervening items, and the long condition had 9-12 intervening items. Each study block had a corresponding recognition memory test block that included one presentation of each study item and an equal number of lures that matched the study items for gender. In order to maintain incidental learning and to eliminate recency effects, we presented all 8 study blocks followed by all 8 test blocks. Study and test blocks each included

48 items. Participants were allowed (but not required) to take a short break in between blocks. After the experiment blocks, each participant also completed the operation span (Ospan) working memory task (Turner & Engle, 1989; automated version: Unsworth, 2005). The revised automated version included strings of 3-7 letters that were interspersed between simple math problems of the form “3+4-2=?” Participants completed 15 trials of the Ospan task, such that they were asked to recall each of 15 strings of letters.

Procedure

Study Phase: Instructions and stimuli were presented on individual monitors. Participants viewed one face at a time for a fixed duration of 1500 ms. During the time the face was on the screen, participants were asked to make a male/female judgment by pressing the “J” key on the keyboard to indicate “male,” or the “K” key to indicate “female.” This judgment task was meant to ensure that participants were paying attention to the stimuli, at least enough to decide if the faces were male or female. Each of the 8 study blocks contained pseudo-random mixing of all four spacing conditions (massed, short, medium, and long). Each face was presented twice within a single study block, such that there were 20 unique stimuli and 40 total stimuli per block. A trial consisted of a fixation cross appearing at the center of the screen for 200 ms. 100-300 ms (mean = 200 ms) after the fixation cross, the first face of the repetition pair appeared on the computer screen for 2000 ms. This was followed by a random jitter that lasted between 100 and 300 ms (mean = 200 ms). Each trial lasted about 1900 ms seconds.

Test Phase: Each study block had a corresponding test block that included all 20 unique items from the study block and 20 lures that were not presented anywhere else in the experiment. The participants were tested for their recognition memory with a task that asked them to press the “J” key on their keyboards if the face was old, or the “K” key if it was new, indicating that they had or had not seen it before. Each stimulus was presented one at a time for a maximum of 2000 ms--after the participants responded with a key press, the face was removed from the screen. With 8 study blocks and 8 test blocks, the task took approximately 23 minutes to complete.

Ospan Task: Prior to the Ospan, the experimenter lead the participants through three practice tasks, which were a letter recall task, a math task, and a sample Ospan trial. In the Ospan task, participants were shown a mathematical expression of the form “3+4-2=?” and pressed “J” or “K” to indicate they had solved it. When a single number appeared, they pressed “J” to indicate that it was the answer to the expression, or “K” to indicate that it was not. If a person took longer than their average time from the math practice plus 2.5 SD (Unsworth et al., 2005), then they were forced to move on and that problem was counted as an error. Once they made their responses to the math problem, a single letter appeared on the screen. When the letter was removed from the screen, a new mathematical expression appeared. This progression of events is illustrated in **Figure B** below.

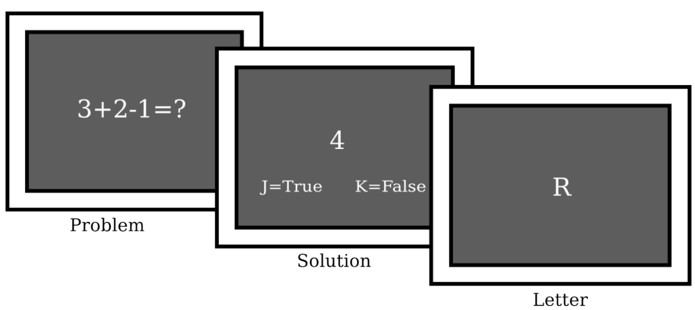

Figure B: Illustration of the progression of events during the Ospan task

After a string of 3-7 letters had been presented, a group of question marks appeared and participants had to recall the entire string. In the hope that participants were not ignoring the math task and simply focusing on remembering the letters, we instructed them to strive for at least 85% accuracy (Unsworth et al., 2005). At the end of each of the 15 trials, a screen appeared that told participants how many mistakes they made on the math task (i.e. “You made 2 math error(s) for this set of trials”) and how many letters they had correctly recalled (i.e. “You recalled 4 letters correctly out of 5”). For letter recall, the program used strict order scoring so that they had to recall the correct letters in their correct serial positions. For example, if the actual string was Y-F-T-J-H and the participant recalled Y-F-T-H-J, they would have scored 3 out of 5 for that string. The task consisted of 15 trials that contained 3 strings of each of the 5 possible lengths, and the string lengths were randomly intermixed. Hence, there was a total of 75 letters and 75 math problems, and the entire task took approximately 20-25 minutes to complete.

Ospan Scoring: Ospan scores were calculated using an absolute scoring method, such that the final score was the sum of the perfectly recalled letter strings. For example, if a person correctly recalled 3 letters in a set of 5, 4 letters in a set of 4, and 6 letters in a set of 6, their score would be 10 (0 + 4 + 6). We imposed a math accuracy criterion of 85% to be sure that the people whose scores we analyzed were putting effort into the math task and not simply devoting all of their mental resources to remembering the letter strings (Unsworth et al., 2005).

Method: Experiment 2

The second study aimed to investigate the interactions among item familiarity, the spacing effect, lag effects, and individual differences in working memory capacity. The experiment consisted of study blocks that were followed by recognition test blocks consisting of all study items and matching lures. We used an equal number of famous faces (current and well-known celebrities including actors, athletes, musicians, politicians, and television personalities; see the Appendix for a list of the famous faces) and non-famous faces. Just like in Experiment 1, repetitions were presented in the massed condition or in one of three spaced conditions, and all conditions were presented in an interleaved manner. The face recognition task was followed by the Ospan task using the same procedures that were employed in Experiment 1.

Participants

78 volunteers were recruited from the Research Experience Program (REP) for Introductory Psychology 1100 students at Ohio State University. All participants had normal or corrected vision, fluency in English, and were at least 18 years of age (M=18.8, SD=1.3). Participants were compensated for their one-hour time commitment with partial course credit. In concordance with standard Institutional Review Board (IRB) protocol, each participant provided consent before beginning the experiment.

Design and Materials

Across all study and test blocks, a total of 384 unique 480 x 480-pixel images were presented. Half of the images (192) were famous faces and the other half of the images were 18 to 50-year-old unfamiliar faces obtained from the database compiled by Minear and Park (2004). Equal numbers of male and female faces were presented from each of the famous and non-famous pools. In an attempt to observe a larger spread of results from the spacing manipulation than we saw in Experiment 1, we updated the lag ranges as follows: the massed condition had 0 intervening items, the short condition had 3-4 intervening items, the medium condition had 7-9 intervening items, and the long condition had 15-18 intervening items. Each study block had a corresponding recognition memory test block that included one presentation of each study item and an equal number of lures that matched the study items for gender and familiarity (famous vs. non-famous). Because overall recognition performance hovered around chance in Experiment 1, we alternated two study blocks and two test blocks until all 16 blocks had been presented instead of massing the study and test blocks. All blocks included 48 items, and participants were allowed (but not required) to take a short break in between blocks.

Procedure

Procedures for Experiment 2 were nearly identical to those of Experiment 1. Participants were instructed to make a male/female judgment via key presses on their keyboards when each face appeared on the screen. Relative to Experiment 1, we increased the presentation duration from 1500 ms to 2000 ms to give participants more time to adapt to each face (see Jacques et al., 2007 for extended explanation of differences in stimulus presentation times). Each of the 8 study blocks contained pseudo-random mixing of all four spacing conditions (massed, short, medium, and long), and random mixing of familiarity conditions (famous vs. non-famous). Each face was presented twice within a single study block, such that there were 24 unique stimuli and 48 total stimuli per block. Each trial lasted for about 2400 ms, and consisted of a fixation cross (200 ms), a single face (2000 ms), and a random jitter (mean = 200 ms). As aforementioned, each pair of study blocks was followed by a corresponding pair of test blocks (*study--study--test--test--study--study--test--test,* etc). Test blocks contained 24 items from its corresponding study block and 24 unique lures. Participants had 2000 ms to respond with a key press as to whether each item was old or new. With 8 study blocks and 8 test blocks, the task took approximately 30 minutes to complete. The Ospan task followed the completion of the main experiment task using the same procedures as Experiment 1.

Results

These data were analyzed with R (R Development Core Team, 2009) and the R packages *lme4* (Bates & Maechler, 2009) and *languageR* (Baayen, 2009; cf. Baayen, 2008) via linear mixed effects models. In order to isolate experimental effects from the variables associated with individual differences in participants, the lmer analysis treats genetic, developmental, social, and environmental factors as random effects (Baayan, Davidson, & Bates, 2008). In addition, differences among the stimuli that may lead to one stimulus being generally more recognizable than the others within the particular stimulus pool (famous vs. non-famous) are treated as random effects as well. Because there was a serious potential for subject effects and item effects to be an issue when analyzing data that depended upon participants' abilities to recognize faces, it was important that we remove these random effects via the linear mixed effects model.

Performance Analysis: Experiment 1

Of the 62 total participants, 13 were excluded from analysis because they did not complete one of the two tasks or their performance on the math task of the Ospan did not exceed 85% accuracy, indicating that they were not suitably distracted from remembering the strings of letters (Unsworth et al., 2005). One participant was excluded for earning a score of 0 on the Ospan task. This left data from 48 subjects (32 females) to be analyzed. Our analyses used a linear mixed effects model, which functions by first accounting for all fixed variables that were observed in Experiment 1, and which are specified within the following equation:

```
m.o = glmer(correct ~ condition * scale(oscore) * scale(study_rt_diff) +
      (1|subject) + (1|stim_name))
```

In plain language, this means that we set the number of correctly recognized target items in the test phase ("correct") as the dependent variable. The operator (~) is read as "is a function of," so the terms to the right represent the independent variables (Baayen, Davidson & Bates, 2008). "Condition" refers to the four spacing conditions: massed, short, medium, and long. "Scale(oscore)" is a continuous variable that includes the normalized Ospan scores for each participant. The term "scale(study_rt_diff)" refers to differences in reaction times to making the male/female judgment during the study phase. Judgment reaction times for the second presentation of each face were subtracted from those of the first presentations. Values for this variable increase as a function of recognition strength due to priming. The random effects for subjects are specified with (1|subject), which adjusts the intercept of the model as a condition of subject groupings. Similarly, random effects for items are specified with (1|stim_name), which adjusts the intercept as a condition of item groupings.

We performed a likelihood ratio test to compare the fit of the full model (m.o) to the fit of the model that did not account for reaction times during the study phase (m.o.nosrt). In R, the likelihood ratio test is carried out with an ANOVA (analysis of variance). We found that the difference was not significant ($p = 0.855$). Therefore, accounting for judgment reaction times during the study phase did not lead to a better fit to the data. O-score (Ospan score) also did not significantly affect the fit of the model ($p = 0.517$), so the most parsimonious model was:

```
m.o.nosrt.noos = glmer(correct ~ condition +
      (1|subject) + (1|stim_name))
```

When performance in the massed condition was set as the intercept of the model, p-values for performance in all spacing conditions were significant such that $p < 0.001$ (massed (intercept): $M = -0.022$, $SE = 0.113$, $p = 0.847$; short: $M = 0.340$, $SE = 0.065$; medium: $M = 0.327$, $SE = 0.065$; long: $M = 0.295$, $SE = 0.065$). **Figure C** on the next page summarizes the effect of repetition spacing on recognition memory performance in Experiment 1.

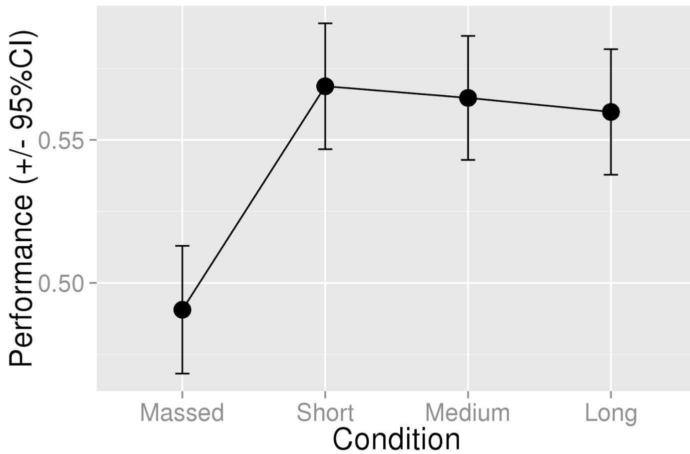

Figure C: Overall recognition performance in Experiment 1 within spacing conditions with a confidence interval of 95%

Priming Analysis: Experiment 1

To see if our performance results were linked to priming effects, our next analysis looked at differences in male/female judgment reaction times during the study phase as the dependent variable. In other words, we wanted to see which independent variables served as predictors of priming effects that occurred during study. Because reaction times reflect processing speed, these priming effects are considered to be a correlate of repetition attenuation. We began our analysis with the following full priming model:

```
m.prime = lmer (study_rt_diff ~ condition * scale(oscore)
      + (1|subject) + (1|stim_name))
```

Using a likelihood ratio test, we found that O-score did not account for a significant amount of variance, but it did trend toward significance ($p = 0.173$). The full model showed a main effect of spacing with the massed condition showing the most priming. There was also a hint of an interaction between condition and O-score.

In these particular analyses, we obtained probability values via Markov chain Monte Carlo (MCMC) methods. Because the variable "study_rt_diff" is calculated by subtracting two reaction times, fitting the above linear mixed effects model returns *t-statistics*, which result from determining whether or not two sets of data are significantly different from one another. At its most basic level, MCMC methods are used to simulate a random sampling of a particular process in order to derive a probability distribution of possible outcomes. In other words, MCMC can take the t-statistics from our reaction time difference calculations and transform them into probability (p) values (Baayen et al., 2008). When priming effects in the massed spacing condition were set as the intercept, p-values were significant in all spacing conditions such that $p < 0.001$ in short, medium, and long (massed (intercept): $t = 27.138$; short: $t = -15.833$; medium: $t = -17.983$; long: $t = -17.920$). **Figure D** below shows the effect of spacing on study phase priming. There is a mild interaction of O-score within the medium and long conditions ($t = 2.100$). Because this interaction is only significant when we isolate the medium and long conditions from the remainder of the data, however, this result will not be discussed further.

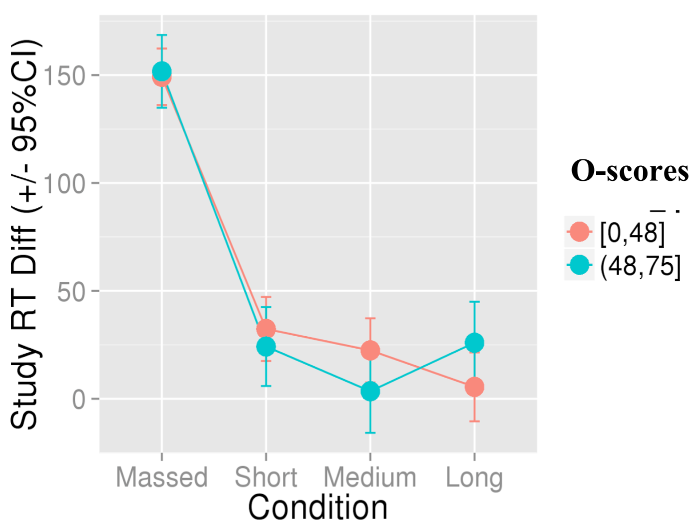

Figure D: Reaction time differences between first and second item presentations during study for high (range from 48 to 75) and low (range from 0 to 48) Ospan scores in Experiment 1 across spacing conditions with a confidence interval of 95%

Performance Analysis: Experiment 2

We excluded 11 of the total 78 participants from analysis because they did not meet the 85% math task accuracy criterion (Unsworth et al., 2005). Data from 67 subjects (37 females) remained to be analyzed using a linear mixed effects model. Our primary model for overall recognition performance for correctly-recognized items accounted for spacing condition, scaled O-scores, stimulus familiarity (famous vs. non famous), and reaction time differences at study with random effects of subjects and items. Likelihood ratio tests revealed that differences in reaction times at study did not significantly affect how well the model fit the data ($p = 0.888$), but Ospan scores did have a significant effect ($p = 0.038$). It should be noted that the most parsimonious model in the analysis of Experiment 1 did not account for Ospan scores, indicating that working memory capacity plays a role in these data specifically when interacting with differences in stimulus familiarity. For the present performance analysis, the final model was as follows:

```
m.o.nosrt = glmer(correct ~ condition * scale(oscore) * stim_familiarity +
      (1|subject) + (1|stim_name))
```

The model showed a main effect of spacing, such that p-values for all three spaced conditions were less than 0.0001 when performance in the massed condition was set as the intercept (massed (interecpt) = 0.115; short = 0.106; medium = 0.106; long = 0.106). There was also a significant main effect of stimulus familiarity ($p < 0.001$). **Figure E** on the next page illustrates these main effects.

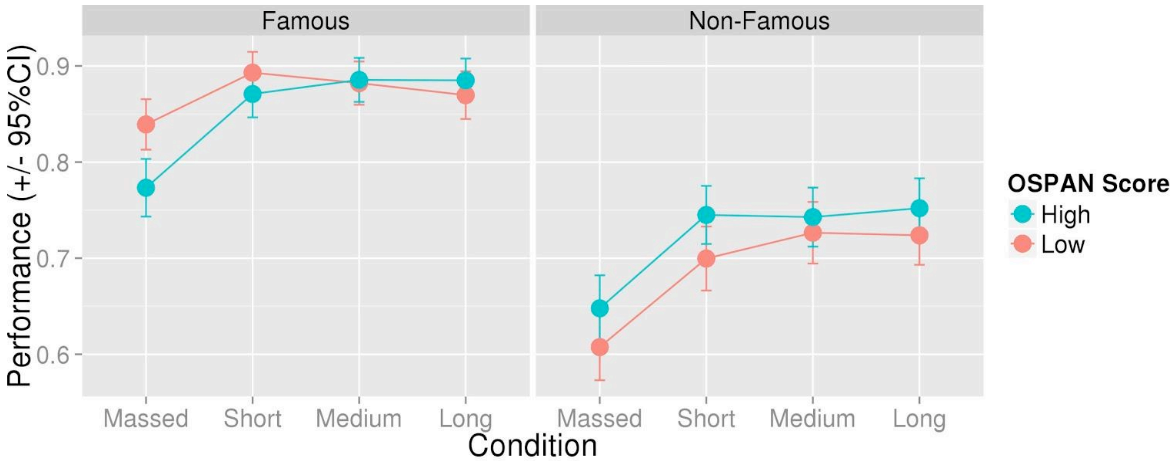

Figure E: Overall recognition performance at test in Experiment 2 split by spacing condition and stimulus familiarity with a confidence interval of 95%; high Ospan scores ranged from 48 to 75 and low scores ranged from 0 to 48

The results indicate that participants were better overall at recognizing famous faces compared to non-famous faces, and they were better at recognizing faces that were repeated under spaced conditions in comparison to massed conditions. Additionally, we found a significant interaction between Ospan scores and stimulus familiarity at test ($p = 0.006$). There was a trending effect of O-score on performance ($p = 0.056$) and a trending interaction effect of O-score and spacing when the long condition was compared to massed ($p = 0.075$).

Priming Analysis: Experiment 2

Our next analysis used differences in reaction times during study as the independent variable. A likelihood ratio test found that O-score did not account for a significant amount of the variance in priming ($p = 0.801$), but condition and stimulus familiarity did. Therefore, our model for the priming analysis was:

```
m.prime = lmer(study_rt_diff ~ condition * scale(oscore) *
stim_familiarity +
      (1|subject) + (1|stim_name))
```

Using MCMC methods to obtain p values, we found a main effect of condition with massed showing the most priming (massed (intercept): $t = 24.798$, $p < 0.001$; short: $t = -13.197$, $p < 0.001$; medium: $t = -13.891$, $p < 0.001$; long: $t = -14.349$, $p < 0.001$). There are also interaction effects of stimulus familiarity and condition when we compare massed to medium ($t = 2.059$, $p = 0.039$) and massed to short ($t = -1.921$, $p = 0.055$). **Figure F** below illustrates these effects, and they will be further investigated in the Discussion section.

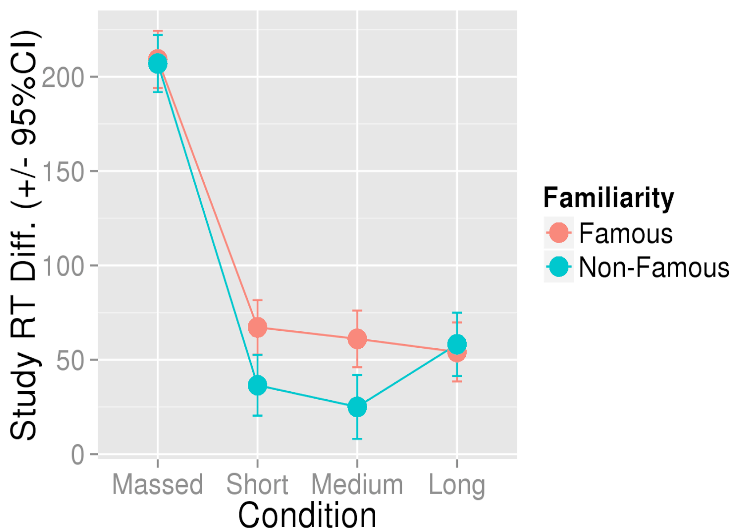

Figure F: Reaction time differences between first and second item presentations during study for famous and non-famous faces in Experiment 2 within spacing conditions with a confidence interval of 95%

Discussion

In Experiment 1, we observed the expected main effect of spacing on recognition memory performance--people more accurately remembered items presented in spaced conditions in comparison to massed conditions. These results were consistent with the findings of Xue et al. (2011), which showed improved recognition accuracy for items that were each repeated 4 times with a mean lag of 12 relative to repeated items that were massed. Our manipulations of spacing lag (short = 2-3 intervening items, medium = 5-7 intervening items, long = 9-12 intervening items), however, did not yield significant differences in subsequent recognition accuracy or in priming during study. This indicates that any amount of spacing at study--whether there were 2 or 12 items in between item repetitions--showed an equal benefit to recognition memory at study and at test in comparison to items that were presented in a massed condition. This is not necessarily surprising because our predictions about performance variations in the separate spacing conditions pertained specifically to differences in Ospan scores. We expected that in the 'long' spacing condition, participants with larger working memory spans--and consequently more control over executive attention (Engle, 2002)--would exhibit better recognition performance than participants with smaller working memory capacities.

In Experiment 1, however, our linear effects model analysis indicated that Ospan scores did not account for a significant amount of the variance in recognition performance. According to our predictions, observing differences in performance within conditions as a function of O-scores would rely upon capturing the moment at which individuals began to forget the first item in the repetition pair by the time the second item in the pair was presented. Given that additional recognition processes are involved in creating a new perceptual representation of an unfamiliar stimulus, then recognizing the second presentation of an item during the study phase is critical for the item being recognized at test. However, it appears that even an interrepetition lag range of 12 is not enough to observe forgetting and resulting drops in recognition performance, regardless of working memory capacity.

The maximum lag interval increased from 9-12 items to 15-18 items from Experiment 1 to Experiment 2, but we believe this was still not enough for forgetting to occur even for people with low working memory capacities--since each trial lasted about 2400 ms, the maximum amount of time in between the first and second item presentations would have been only 43 seconds. In a study on forgetting conducted by Ferris and colleagues, the researchers presented participants with images of faces for 5 seconds each and repeated the images at lags of 0.5, 1, 2, 4 or 40 minutes (Ferris, Crook, Clark, McCarthy & Rae, 1980). Using these manipulations, the researchers observed significant decreases in recognition performance at the time of the second presentation of the items within each subsequent condition. In future studies, we would extend both the adaptation time and the amount of time in between repetitions in an effort to observe repetition effects in action, and see how Ospan scores relate to performance curves for recognizing familiar and unfamiliar items.

Although Ospan scores did not significantly account for variation in performance in Experiment 1, it did account for more variation than a null model in Experiment 2. One possible reason for this discrepancy relates to how working memory stores unfamiliar stimuli in comparison to familiar. There is evidence that subverbal rehearsal is used to hold lists of digits, words, and letters in working memory (Baddeley, 1986; Baddeley & Gathercole, 1998); however, there is no evidence that subverbal rehearsal occurs for unfamiliar faces. Labels are assigned to images so that they can be verbally rehearsed (Lutz & Scheirer, 1974), but this does not necessarily occur for unfamiliar images within a category. For example, if a list contains images of a duck, a chair, and an apple, individuals will presumably code the images verbally in addition to coding them visually. When looking at successive images of unfamiliar faces, the lack of pre-existing associations for each individual face prevents distinguishable labeling that would enable rehearsal. While successful subverbal rehearsal is critical for high performance on the Ospan task, it is inappropriate for remembering unfamiliar faces--this is a potential explanation of why there was no effect of O-score on performance in Experiment 1. Experiment 2, however, involved familiar faces that participants would presumably be able to subverbally label, assuming they knew the names of the celebrities presented. Hence, O-scores made a difference only when we included the familiar stimuli. Labeling of famous faces may also contribute to why memory performance was so much better for famous faces compared to non-famous faces. Paivio's dual coding theory posits that familiar images are coded both verbally and pictorially (1967). Access to pre-existing representations combined with additional coding for familiar items theoretically leads to improved recognition memory compared to only pictorial coding.

In Experiment 2, we saw a peculiar interaction of working memory capacity and stimulus familiarity that appears to be driven by the fact that at shorter lags, people with low Ospan scores are better than people with high Ospan scores at recognizing famous faces at test (see **Figure E**). Though it is perplexing that people with small WM capacities outperformed people with large WM capacities on a test of recognition memory, this result is actually consistent with findings by McCutcheon, Ashe, Houran and Maltby (2003). Their study successfully linked celebrity-worship to cognitive deficits, including deficits in intelligence. Since the Ospan and other measures of WM capacity are correlated with measures of general fluid intelligence (Conway, Cowan, Bunting, Therriault & Minkoff, 2002), it appears people with low O-scores are more accurate at recognizing famous faces at test because they pay more attention to them at study compared to people with high O-scores. In other words, the difference in performance may arise because celebrities are more meaningful to people with lower fluid intelligence. The trending interaction effect of O-score and spacing when the long condition was compared to massed seems to be driven by the same result. This explanation is post hoc, of course, and should be investigated with further studies including multiple measures of intelligence and a focus on how intelligence interacts with memory for celebrities.

We performed analyses of priming effects to examine the bases of the performance results discussed previously. In both Experiment 1 and Experiment 2, people were faster at making the male/female judgment of the second face in the repetition pair relative to first face in the pair when the images were massed rather than spaced. Additional processing is not required to make a male/female judgment for the second face presented in a massed condition because participants would have made the exact same judgment of the exact same face immediately prior. When intervening items are present, however, participants must process the stimulus again in order to make a judgment. Priming results for Experiment 1 were consistent with results from a study by Bentin & Moscovitch (1988), which reported significant repetition effects for unfamiliar faces presented with a lag of 0 (the other possible lag lengths were 4 and 15). The authors posited that the memory traces tied to the first presentation of the unfamiliar faces were too weak for priming effects to occur at longer lags. This explanation seems plausible, particularly when we consider our results of changes in reaction times during study in Experiment 2: we observed a significant difference in priming effect in the medium condition and an effect that trended toward significance in the short condition when we compared the data for famous and non-famous faces. The familiarity of the famous faces allowed them to be processed more deeply so that their memory traces were consequently stronger. Hence, participants had faster reaction times for the second presentations of faces during study even when other stimuli intervened between repetitions.

The aforementioned results and implications invite questions for further study. In Experiment 2, we found a main effect of stimulus familiarity on recognition performance that mirrored the findings of Henson and colleagues in their 2002 study. The experimenters related their results to fMRI data that showed neural attenuation for repeated images of famous faces and neural enhancement for repeated images of non-famous faces. In Xue et al.'s 2011 study that used only unfamiliar face stimuli, however, items that were presented in a massed condition resulted in more repetition attenuation effects and lower recognition memory performance at test in comparison to items that were presented in a spaced condition. The present study sought to reconcile these results by manipulating repetition priming via lag conditions and comparing recognition performance to familiar and unfamiliar stimuli while accounting for differences in working memory capacity. Without capturing the boundary between recognition and forgetting during study, however, we are unable to draw conclusions about repetition attenuation. Future studies should extend the 'long' spacing condition to the point that participants begin to forget the first item in a repetition pair before the second item in the pair is presented. In addition, future work will investigate the neural correlates of lag via repetition attenuation.

References

Baayen, R.H. (2008). Analyzing Linguistic Data: A Practical Introduction to Statistics Using R. Cambridge: Cambridge University Press.

Baayen, R.H., Davidson, D.J., Bates, D.M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. Journal of Memory and Language, 59, 390-412.

Baayen, R. H. (2009). languageR: Data sets and functions with "Analyzing Linguistic Data: A practical introduction to statistics". R package version 0.955.

Baddeley, A. D. (1986). *Working Memory*. Oxford: Oxford University Press.

Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological review*, *105*(1), 158.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. *The psychology of learning and motivation*, *8*, 47-89.

Bates, D.M. & Maechler, M. (2009). lme4: Linear mixed-effects models using S4 classes. R package version 0.999375-32.

Bentin, S., & Moscovitch, M. (1988). The time course of repetition effects for words and unfamiliar faces. *Journal of Experimental Psychology General*, *117*, 148-148.

Brewer, J. B., Zhao, Z., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. (1998). Making memories: brain activity that predicts how well visual experience will be remembered. *Science*, *281*(5380), 1185-1187

Bruce, V., & Young, A. (1986). Understanding face recognition. *British journal of psychology*, *77*(3), 305-327.

Callan, D. E., & Schweighofer, N. (2010). Neural correlates of the spacing effect in explicit verbal semantic encoding support the deficient‐processing theory. *Human brain mapping*, *31*(4), 645-659.

Case, R., Kurland, D. M., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. *Journal of experimental child psychology*,*33*(3), 386-404.

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological bulletin*, *132*(3), 354.

Conway, A. R., Cowan, N., Bunting, M. F., Therriault, D. J., & Minkoff, S. R. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence*, *30*(2), 163-183.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of verbal learning and verbal behavior*, *19*(4), 450-466.

Donovan, J. J., & Radosevich, D. J. (1999). A meta-analytic review of the distribution of practice effect: Now you see it, now you don't. *Journal of Applied Psychology*, *84*, 795-805.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. (1999). Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. *Journal of Experimental Psychology: General*, *128*(3), 309.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current directions in psychological science*, *11*(1), 19-23.

Ferris, S. H., Crook, T., Clark, E., McCarthy, M., & Rae, D. (1980). Facial recognition memory deficits in normal aging and senile dementia. *Journal of Gerontology*, *35*(5), 707-714.

Greene, R. L. (1989). Spacing effects in memory: Evidence for a two-process account. *Journal of experimental psychology. Learning, memory, and cognition*,*15*(3), 371-377.

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends in cognitive sciences*, *10*(1), 14-23.

Henson, R. N. A., Price, C. J., Rugg, M. D., Turner, R., & Friston, K. J. (2002). Detecting latency differences in event-related BOLD responses: application to words versus nonwords and initial versus repeated face presentations. *Neuroimage*, *15*(1), 83-97.

Henson, R., Shallice, T., & Dolan, R. (2000). Neuroimaging evidence for dissociable forms of repetition priming. *Science*, *287*(5456), 1269-1272.

Jacques, C., d'Arripe, O., & Rossion, B. (2007). The time course of the inversion effect during individual face discrimination. *Journal of Vision, 7,* 1-9.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *The Journal of Neuroscience*, *17*(11), 4302-4311.

Lutz, W. J., & Scheirer, C. J. (1974). Coding processes for pictures and words. *Journal of Verbal Learning and Verbal Behavior*, *13*(3), 316-320.

Maddox, G. B., Balota, D. A., Coane, J. H., & Duchek, J. M. (2011). The role of forgetting rate in producing a benefit of expanded over equal spaced retrieval in young and older adults. *Psychology and aging*, *26*(3), 661.

McCabe, D. P., Roediger III, H. L., McDaniel, M. A., Balota, D. A., & Hambrick, D. Z. (2010). The relationship between working memory capacity and executive functioning: evidence for a common executive attention construct. *Neuropsychology*, *24*(2), 222.

McCutcheon, L. E., Ashe, D. D., Houran, J., & Maltby, J. (2003). A cognitive profile of individuals who tend to worship celebrities. *The Journal of Psychology*, *137*(4), 309-322.

Minear, M. & Park, D.C.(2004). A lifespan database of adult facial stimuli. Behavior Research Methods, Instruments, & Computers. 36, 630-633.

Paivio, A. (1967). Paired-associate learning and free recall of nouns as a function of concreteness, specificity, imagery, and meaningfulness. *Psychological Reports*, *20*(1), 239-245.

Park, D. C., Smith, A. D., Lautenschlager, G., Earles, J. L., Frieske, D., Zwahr, M., & Gaines, C. L. (1996). Mediators of long-term memory performance across the life span. *Psychology and aging*, *11*(4), 621.

R Development Core Team (2009) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Russo, R., Parkin, A. J., Taylor, S. R., & Wilks, J. (1998). Revising current two-process accounts of spacing effects in memory. *Journal of Experimental Psychology Learning Memory and Cognition*, *24*, 161-172.

Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of Experimental Psychology: General*, *125*(1), 4.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent?. *Journal of Memory and Language*, *28*(2), 127-154.

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior research methods*,*37*(3), 498-505.

Wagner, A. D., Koutstaal, W., Maril, A., Schacter, D. L., & Buckner, R. L. (2000). Task-specific repetition priming in left inferior prefrontal cortex. *Cerebral Cortex*, *10*(12), 1176-1184.

Wagner, A. D., Schacter, D. L., Rotte, M., Koutstaal, W., Maril, A., Dale, A. M., Rosen, B.R., & Buckner, R. L. (1998). Building memories: remembering and forgetting of verbal experiences as predicted by brain activity. *Science*, *281*(5380), 1188-1191.

Xue, G., Mei, L., Chen, C., Lu, Z. L., Poldrack, R., & Dong, Q. (2011). Spaced learning enhances subsequent recognition memory by reducing neural repetition suppression. *Journal of cognitive neuroscience*, *23*(7), 1624-1633.

Appendix
List of famous face stimuli used in Experiment 2

Al Roker
Alec Baldwin
Allen Iverson
Anderson Cooper
Andre Agassi
Arnold Schwarzenegger
Ashton Kutcher
Barack Obama
Ben Stiller
Bill Cosby
Bill Gates
Billy Crystal
Bob Saget
Bono
Brad Pitt
Christopher Reeve
Christopher Walken
Colin Powell
Conan O'Brian
Dalai Lama
Daniel Craig
Daniel Radcliffe
Denzel Washington
Donald Trump
Dustin Hoffman
Ed Helms
Eddie Murphy
Elvis Presley
Eminem
George Clooney
George W. Bush
Harrison Ford
Heath Ledger
Hugh Grant
Hugh Laurie
Hulk Hogan
Jack Black
Jack Nicholson
Jake Gyllenhaal
James Marsden
Jay Leno
Jerry Seinfeld
Jim Carrey
Joe Biden
John F. Kennedy
John Kerry
John Lennon
John Mayer
John McCain
John Travolta
Johnny Depp
John Stweart
Justin Timberlake
Kanye West
Keanu Reeves
Kurt Cobain
Lance Armstrong
Larry King
LeBron James
Leonardo DiCaprio
Martin Luther King Jr.
Matt Lauer
Matthew Perry
Mel Gibson
Michael Jackson
Michael Phelps
Morgan Freeman
Nelson Mandela
Nicolas Cage
Orlando Bloom
Owen Wilson
Patrick Dempsey
Patrick Stewart
Pierce Brosnan
Ray Charles
Ray Romano
Regis Philbin
Richard Nixon
Richard Simmons
Robert DeNiro
Ronald Reagan
Russell Crowe
Shia Lebeouf
Simon Cowell
Snoop Dogg
Stephen Colbert
Steve Carell
Sylvester Stallone
Tiger Woods
Tim McGraw
Tobey Maguire
Tom Cruise
Tom Hanks
Tracy Morgan
Usher
Vince Vaughn
Zac Efron
Zach Braff

Adele
Alyson Hannigan
America Fererra
Amy Poehler
Angelina Jolie
Anne Hathaway
Ashley Tisdale
Audrey Hepburn
Barbara Walters
Betty White
Beyonce Knowles
Blake Lively
Britney Spears
Cameron Diaz
Carrie Underwood
Catherine Zeta Jones
Charlize Theron
Chelsea Handler
Cher
Christina Aguilera
Condoleezza Rice
Demi Moore
Dolly Parton
Drew Barrymore
Ellen Degeneres
Emma Stone
Emma Watson
Eva Longoria
Faith Hill
Fergie
Gwen Stefani
Halle Berry
Heidi Klum
Hillary Duff
Hilary Swank
Hillary Clinton
Jenna Fischer
Jennifer Aniston
Jennifer Lopez
Jessica Alba
Jessica Simpson
Julia Roberts
Julie Andrews
Kate Hudson
Kate Middleton
Katie Couric
Katie Holmes
Katy Perry
Keira Knightley
Kelly Clarkson
Kesha
Khloe Kardashian
Kim Kardashian
Kirsten Dunst
Kristen Stewart
Lady Gaga
Lindsay Lohan
Liv Tyler
Madonna
Mariah Carey
Marilyn Monroe
Martha Stewart
Megan Fox
Melissa McCarthy
Meryl Streep
Michelle Obama
Miley Cyrus
Natalie Portman
Nicki Minaj
Nicole Kidman
Oprah Winfrey
Paris Hilton
Paula Deen
Penelope Cruz
Pink
Princess Diana
Rachael Ray
Rebel Wilson
Reese Witherspoon
Renee Zellweger
Rihanna
Rosie O'Donnell
Salma Hayek
Sandra Bullock
Sarah Michelle Gellar
Sarah Palin
Scarlett Johansson
Selena Gomez
Serena Williams
Shakira
Shawn Johnson
Sofia Vergara
Taylor Swift
Tina Fey
Tyra Banks
Uma Thurman
Vanessa Hudgens
Whoopi Goldberg
Zooey Deschanel